Columbia University

Department of Economics
Discussion Paper Series

Default Tips

Kareem Haggag, Giovanni Paci

Discussion Paper No.: 1213-21

Department of Economics
Columbia University
New York, NY 10027

March 2013

Default Tips

Kareem Haggag & Giovanni Paci *

March 18, 2013

Abstract We examine the role of defaults in high-frequency, small-scale choices using unique data on over 13 million NYC taxi rides. We exploit a shift in the set of default tip suggestions presented to customers prior to payment, as the base fare changes from below \$15 to above \$15. Using a regression discontinuity design, we show that default suggestions have a large impact on tip amounts. These results are supported by a secondary analysis that uses the quasi-random assignment of customers to different cars to examine default effects on all fares above \$15. Finally, we highlight a potential cost of setting defaults too high, as a higher proportion of customers opt to leave no credit card tip when presented with the higher suggested amounts.

* **Haggag:** University of Chicago, Booth School of Business; kareem.haggag@chicagobooth.edu. **Paci:** Columbia University, Economics Department; gp2282@columbia.edu. We thank Stephan Meier, Devin Pope, Bernard Salanie, and Eric Verhoogen for reading earlier drafts of this paper and providing excellent suggestions that greatly improved its quality. We also thank Colin Camerer, Ray Fisman, Eric Johnson, Supreet Kaur, John List, Chenfei Lu, Suresh Naidu, Adam Osman, Kiki Pop-Eleches, Sebastien Turban and Miguel Urquiola for helpful comments and discussions at various stages of this project.

1 Introduction

The large effects of default options on consumer choices have been documented in various high-stakes, but low-frequency contexts, ranging from organ donation to 401(k) contributions. Because defaults preserve freedom of choice, but nonetheless appear to strongly influence behavior, they have been of great interest to both policy makers and academics (*Nudge*, Thaler & Sunstein 2008). In contrast to the extant literature, we study the effects of defaults on a frequently encountered consumer choice: the decision of how much to tip a service provider.[1] By studying tipping, we demonstrate the ability of defaults to nudge behavior in a decision problem which agents have arguably encountered enough times to learn their optimal responses. In doing so, we also extend the literature by documenting a case in which default effects were exploited by a for-profit industry.

Our study introduces a unique data set that contains fare information for 170 million NYC taxi rides over the calendar year of 2009. Among these rides, we have tip information for the 38 million credit card transactions, from which we use a sample of 13 million rides to study tipping.[2] At the end of each ride, customers who used credit cards were presented with a screen that provided them with the option to either type in a desired tip amount or to press one of three buttons with default tip suggestions. During the period of study, one of the credit card machine companies offered different tip suggestions depending on whether the fare was above or below \$15.[3] For rides under \$15, tip suggestions were \$2, \$3, and \$4, while rides above \$15 were presented with 20%, 25%, and 30% tip suggestions. At the discontinuity, this shift represents an increase in the suggestion categories (low, medium, and high) of approximately \$1, \$0.75, and \$0.50 for rides without tolls, taxes, or surcharges. Importantly, the shift in suggestions did not change the choice set; customers were still free

[1]Though difficult to precisely measure, Azar (2011) estimates total annual tipping in the US food industry alone at \$46 billion, or approximately 0.3% of annual GDP.

[2]Our sample of study consists of the entire universe of credit card transactions during a period of time in which there were no tolls, taxes, or surcharges.

[3]More precisely, the threshold is determined by the "base amount" (the sum of the fare, taxes, tolls, and surcharges); however, in the sample we focus on, the base amount is equivalent to the fare.

to key in any tip amount. Under the assumption that all ride characteristics that affect tips vary smoothly with the base amount, the difference at the discontinuity can be used to identify the causal effect of this particular increase in default suggestions on tipping. We find that this local treatment effect is an increase in tip amounts of approximately \$0.27 - \$0.30, greater than a 10% increase in the average tip at that margin.

Part of the observed default effect at the discontinuity may be attributed to the difficulty of converting between dollars and percentages (Kahneman 2011).[4] To address this concern and to examine the role of default suggestions across a larger range of fares, we present a second econometric strategy. We use the quasi-random assignment of passengers to taxi cabs at LaGuardia airport to compare across credit card machine companies. For rides above \$15, both companies provided percentage defaults; however, one company provided 15%, 20%, and 25% percent, while the other provided defaults of 20%, 25%, and 30% percent. The distribution of tips clearly reflects this shift, and again, we find that higher defaults are associated with higher average tip amounts, controlling for time-invariant driver characteristics.

Having demonstrated the benefits of higher default suggestions on the intensive margin of tipping, we next highlight a potential cost of setting defaults too high. First, in both the regression discontinuity design and the comparison across vendors, we find that the higher default suggestions reduce the probability of leaving a tip that corresponds to one of the tip suggestions (24 and 7.8 percentage point reductions respectively).[5] More striking is the result that rides with the higher tip suggestions are almost twice as likely to receive a zero-valued tip as their competitors (a 2.8 percentage point increase). Such customers may have

[4]On its own, the difficulty of comparing across the measurements does not imply that we should find higher tip amounts for the percentage suggestions. One potential explanation of this pattern would rely on this computational difficulty interacting with a particular type of self-deception. If customers adhere to tip *percentage* norms, then dollar suggestions could result in less generous tips by lowering the cost of self-deception. For example, consider a customer that has a fare of \$13 and adheres to a 25% tipping norm (i.e. a tip of \$3.25). This customer may be able to convince herself that she is adhering to the norm by selecting the \$3 option (rounding in the direction of her self-interest), whereas she could not ignore her deviation from the norm if explicitly presented with the 25% option.

[5]However, we also find that the average manually entered tip amount increases. Thus, it's not clear that those induced to leave a manual tip are leaving lower tips than they would with the lower suggestions.

been penalizing drivers for using tip defaults that are perceived as *unfairly* high.

Finally, we investigate heterogeneity and shed some light on potential mechanisms. Several factors may explain our observed default effects. Customers may be rationally inattentive, failing to compute their preferred tip due to the opportunity cost of time and/or the cognitive costs associated with that computation. Moreover, customers that are unfamiliar with the tipping norm may interpret the defaults as the socially endorsed norm. Both uninformed and informed customers may experience disutility from deviating from these options. To investigate these mechanisms, we exploit the geocoding of our data to merge the pick-up and drop-off location census tracts to the American Community Survey. Ultimately, the data do not provide strong support for one mechanism over the other. Indeed, several attempts at splitting the sample suggest a striking robustness and constancy of the effect across different possible types of customers.

We build upon the broad literature on defaults. Default effects have been demonstrated across a wide variety of consumer choices. Most notably, Madrian and Shea (2001) and Choi et al. (2004) found large effects in retirement savings contributions, with Madrian and Shea (2001) finding a 50% increase in enrollment from switching from an opt-in to automatic enrollment default.[6] In a similarly sparsely encountered consumer choice, Goldstein and Johnson (2003) and Abadie and Gay (2006) used cross-country analyses to suggest that presumed consent policies induce higher organ donation rates than opt-in policies. Johnson et al. (1993) studied a somewhat more frequently encountered type of decision problem, namely (car) insurance plan choice.[7] Downs et al (2009) found that a convenience manipulation that approximates a default influenced food purchase decisions.[8] Our paper contributes by showing that default effects can persist in a similarly habitually encountered consumer

[6] Similarly, Cronqvist and Thaler (2004) found evidence of large default effects in their study of the Swedish social security privatization.

[7] Furthermore, Johnson et al (2002) studied default effects in the decision to accept email marketing.

[8] In their experiment, fast-food restaurant patrons were provided with a free meal upon completion of a survey. Participants were provided with a survey binder followed by a one-page menu featuring either five low-calorie, five high-calorie, or a mix of sandwiches – other options were provided in a pamphlet at the back of the binder. Relative to the mixed-calorie condition, subjects in the low-calorie treatment were 47% more likely to choose a low-calorie sandwich.

choice, using a much larger naturalistic field data set. We also add to the literature by tracing out the response to higher defaults, including its limitations. Beshears et al (2010) similarly study the limits to setting high defaults. They provide a case study of a firm that set the default contribution rate at 12%, a rate much higher than previously studied defaults in this area (2% - 6%) and one that the authors note was sub-optimal for all employees in the sample. They find that roughly 25% of employees remain at this default rate after 12 months of tenure, in comparison to the 60% adherence rate seen at firms in previous studies. In our study, we find that a substantial proportion are induced to opt out of the default when presented with the higher suggestions. We still find a higher average contribution despite this result; however, our analysis also highlights the emergence of a cost (zero-valued tips) that suggests a potential reduction in tips if defaults are set sufficiently high.[9]

Our paper also relates to an emerging literature on active choice; a form of design in which no default option is provided, and instead, consumers are forced to explicitly choose an option. Carroll et al (2009) studied the effects of switching from an opt-in regime to active choice in 401(k) plan enrollment, finding a 28% increase in enrollment under active choice relative to opt-in. The language used in the active choice condition was careful to not favor one choice over the other (I want to enroll vs. I don't want to enroll). In contrast, Keller et al (2011) ran a set of small experiments that tested the relative effectiveness of opt-in, active choice, and enhanced active choice (using language that favored one option). Our setting is similar to enhanced active choice, as all customers were still required to make an active choice – a tip amount had to be selected to complete payment. However, as in the enhanced active choice studies, certain options were favored, in our case by convenience and by the potentially implied endorsement of being set apart. Our analysis uniquely combines clean identification of the effect of enhanced active choice with the naturalism of field data.

[9]This cost of setting defaults too high also relates to the results of a charitable giving field experiment by Karlan, List, and Shafir (2011). In their experiment, matching grants were framed as either $1 for every $1/$3 or $25 for every $25/$75. They found that using higher example amounts resulted in fewer and smaller donations. While this particular effect is open to a number of possible interpretations, it is consistent with a potential backfiring of setting implied appeals too high (even when those appeals are just example amounts that are not convenience advantaged).

The paper proceeds as follows. Section 2 provides background on taxis and tipping and describes the data used. Section 3 presents our regression discontinuity results and robustness tests. Section 4 presents an analysis that compares across credit card machine companies. Section 5 presents results on the cost of setting defaults too high. Section 6 explores heterogeneity and mechanisms. Section 7 concludes.

2 Institutional Context and Data

The data for our study was provided by the Taxi and Limousine Commission (TLC) of New York City. In May 2004, the TLC mandated that all taxi cabs be outfitted with a set of technological improvements, including the electronic collection and transmission of trip data and the introduction of equipment to accept credit cards.[10] These technological improvements also marked the introduction of a system that measured and saved the GPS coordinates of all pick-up and drop-off locations. Though mandated in 2004, the entire taxi fleet was not outfitted with the equipment until 2008.[11] Our data spans the entirety of 2009, covering all rides by licensed Yellow Cab drivers in NYC. Before describing the data, we first present the details on the institutional context.

2.1 Institutional Context

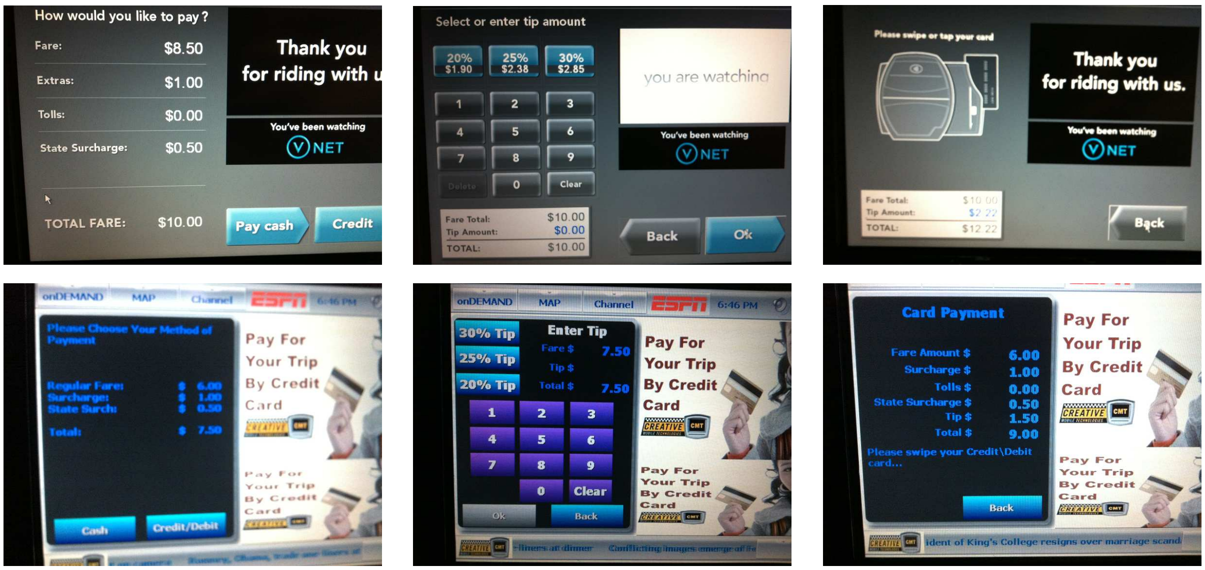
During the period of study, three companies were contracted to provide taxi cabs with credit card machines. The largest two, which we denote as "Vendor" and "Competitor", account for 49% and 45% of observations in the raw data respectively. Each taxi cab was equipped with its own Passenger Information Monitor (PIM) which would display advertisements and other viewing material during the ride. At the end of the ride, the PIM displayed a payment screen. Figure 1 provides an example of a payment screen that would be presented

[10]Source: http://www.nyc.gov/html/tlc/html/industry/taxicab_serv_enh.shtml.

[11]Source: "Despite some grumbling, however, the TLC is moving to install the devices in all cars by August 31." http://www.nysun.com/business/hot-tip-for-cabbies-credit-cards-boost-tips/72783/.

to customers.

Figure 1: Example of a Passenger Information Monitor (PIM) payment screens by the Vendor (Top) and the Competitor (Bottom)

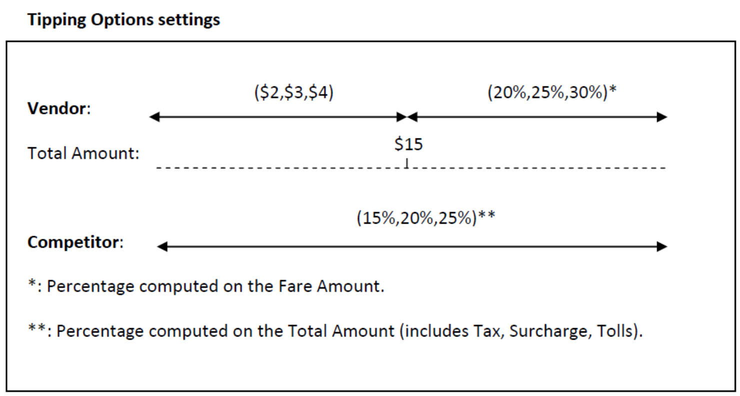


Notes: The top and bottom rows of screens correspond to Vendor and Competitor, respectively. The sequence of payment screens follows from left to right. The authors took these screen-shots from taxi cabs in October 2012. During the period pictured, both the Vendor and Competitor only provide 20%/25%/30% default tips, and both compute these tips on only the fare. In contrast, during the period of our study (2009), the Vendor offered defaults of $2/$3/$4 for amounts less than $15 and defaults of 20%/25%/30% (computed on the fare) for fares greater than $15, while the Competitor offered defaults of 15%/20%/25% (computed on the sum of fare, tax, surcharge, and tolls) for all fares.

Customers were presented with the base amount and had the option of keying in their own tip amount or using one of the suggested tip buttons. Each vendor was allowed discretion over how this page was displayed, and the two companies elected to offer different default buttons during the period we study.

There were two key ways in which the Vendor and Competitor differed with respect to tip suggestions. The Competitor offered three suggestions on all base amounts: 15%, 20%, or 25%. In contrast, the Vendor provided one set of suggestions ($2, $3, or $4) for all base amounts lower than $15, and another set of suggestions (20%, 25%, or 30%) for all base amounts above $15. The second difference is how the percentages were calculated. Though the Vendor used the base amount (i.e. the fare, toll, MTA tax, and surcharge) to determine which set of suggestions to provide, the percentage tips were calculated on only the fare. The Competitor instead calculated percentages on the entire base amount. Thus, if the ride consisted of a \$10 fare and a \$10 toll, a customer that pushed the 20% button with the Vendor equipped cab would be paying \$2 in tip, whereas a customer in a Competitor equipped cab would be paying \$4. These differences are summarized in Figure 2. A third difference that currently exists, but may not have during the study period, is that the Vendor also displays corresponding dollar amounts with the percentage tips, while the Competitor does not display these conversions.

Figure 2: Tip Default Suggestions by Vendor and Competitor



Taxi meters determined the fare through a combination of time and distance measures. The standard city rate (Rate Code 1) charged customers \$2.50 upon entry, and \$0.40 for each additional unit.[12] One unit is defined as either (1) a 60 second interval in which the car is idle or driving less than 6 miles per hour or (2) 0.20 miles when the car is driving 6

[12]Other Rate Codes include:

- Rate Code 2 - Rides to and from JFK - Charged a flat rate of \$45.
- Rate Code 3 - Rides to and from Newark Airport - Charged the standard rate in addition to a \$15 surcharge.
- Rate Code 4 - Rides to Nassau or Westchester county - Charged the standard city rate while in city limits, and double the standard rate while in Nassau or Westchester county.
- Rate Code 5 - Rides outside NYC, excluding Nassau, Westchester, or Newark Airport - Charged flat rate (determined through negotiation between rider and driver).

Source: http://www.nyc.gov/html/tlc/html/passenger/taxicab_rate.shtml.

miles per hour or faster. Fractional amounts are rounded up to the next unit. Riders were also subject to different sets of surcharges depending on the period of the year or the period of the day. We define six pricing regimes accordingly. The first three pricing regimes cover the period of January 1, 2009 to October 31, 2009. Pricing regime one spans 6am to 4pm on Monday - Friday and 6am to 8pm on Saturday and Sunday. Customers were not subject to any surcharges during this first pricing regime. The second regime covers 4pm to 8pm on Monday through Friday. During these peak weekday hours, customers were subject to a surcharge of $1.00. The third regime covers 8pm to 6am on all days, during which customers were subject to a $0.50 nighttime surcharge. Regimes 4 to 6 cover the period of November 1, 2009 to December 31, 2009. These last three regimes are similar to 1-3, except that riders were additionally subject to a $0.50 MTA tax. To maintain comparability on either side of the Vendor default discontinuity, we limit our primary analysis to pricing regime one, which corresponds to a plurality among these regimes ($>$ 34% of rides). We further restricted this sample to rides that did not pass through tolls. During this regime, the largest base amount to the left of the $15 discontinuity was $14.90, and the largest base amount to the right of the discontinuity was $15.30.

2.2 Data Description

Our preliminary data set included 170,896,479 observations. Though the TLC has its own private routine for removing electronic glitches, the provided data set still contained a number of possible electronic errors, including zero-valued distances/durations and surcharges that did not correspond to the appropriate schedule. We took a number of steps to clean the data, which we outline in greater detail in Appendix A. Our largest sample reductions were the removal of Cash payments (approximately three-quarters of the sample), and limiting the sample to the pricing regime under which there were no taxes or surcharges. Our final dataset consists of 13,820,735 rides. For the majority of our estimates, we limit our sample to rides complete on cars equipped by the Vendor (7.26 million) and to fares between $5 and

\$25 (6.24 million observations).

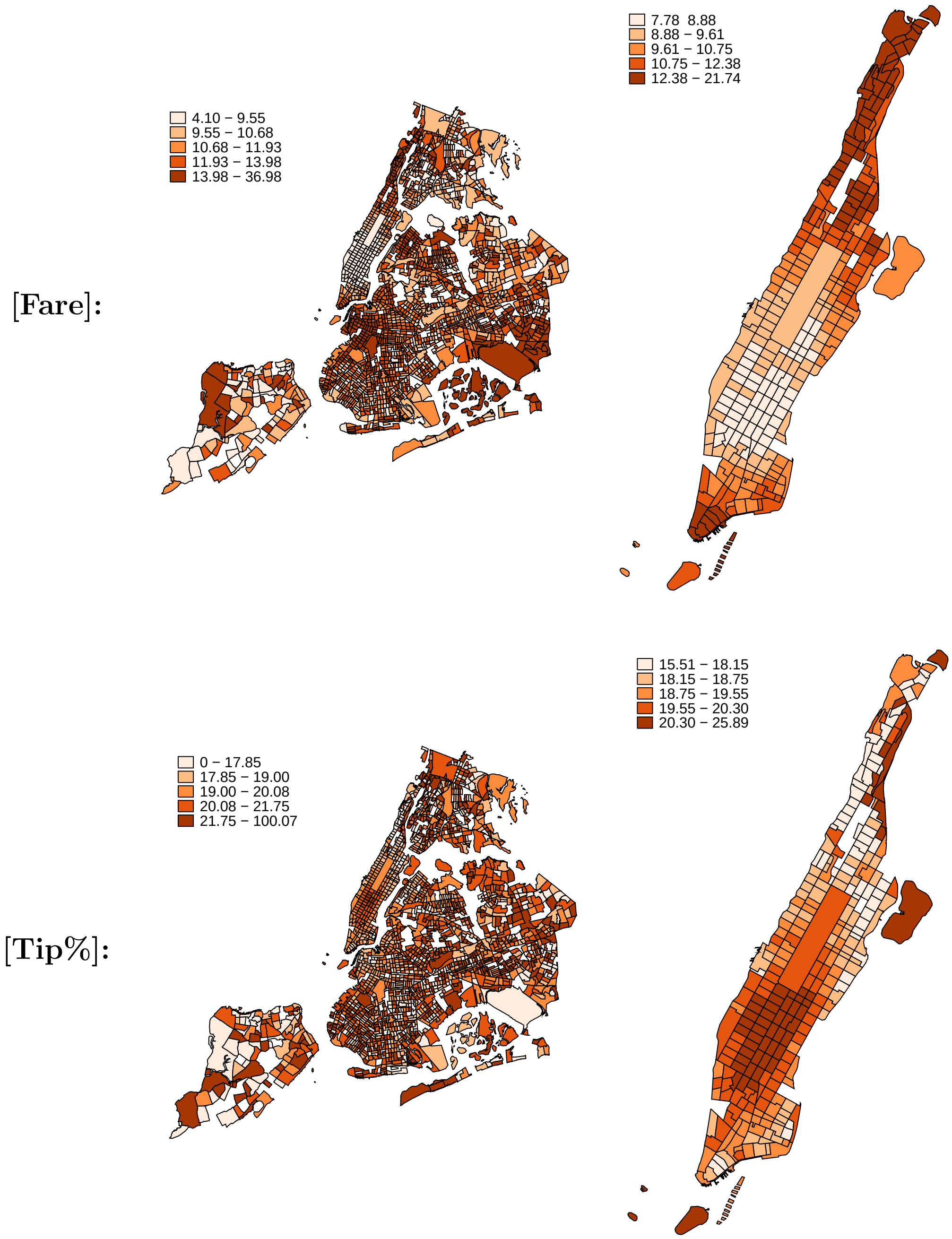
The variables provided in the data are as follows: anonymized driver identifier, car identifier, credit card machine company, payment type, ride duration, ride distance, number of passengers, fare, surcharge, MTA tax, toll amount, and pick-up and drop-off latitude, longitude, and time. Because we do not have an indicator for whether the customer physically selected one of the default suggestion buttons, we needed to create this key variable. To do so, we make the assumption that all tip amounts that correspond to one of the relevant tip suggestions (e.g. \$2/\$3/\$4 for "Vendor" if the base amount is less than \$15) were selected from one of these buttons. We thus make the assumption that customers recognize this congruence and save the time of keying in this amount by pressing a single button.

For the purpose of computing heterogeneous treatment effects, we use data from the American Community Survey's 5 year estimates (2005 - 2009). This dataset provides census tract level summary statistics. We match these statistics to each pickup and drop-off location. To do so, we first assign each GPS coordinate to a census tract using a point-in-polygon operation in PostgreSQL (PostGIS). We then merge each pickup location and each drop-off location with the ACS census tract variables. We focus on one variable in particular: median household income.[13]

To better understand the data, we provide a few descriptive tables and figures. Figure 3 plots the geospatial distribution of fares and tip percentages, splitting the pick-up location census tracts into five quantiles and shading accordingly. Table 1 cross tabulates the pick-up and drop-off boroughs of all trips in our final sample. More than 90% of rides both start and end in Manhattan. Table 2 provides a table of summary statistics by ride. The average fare is \$9.75 and the average tip percentage is 18.13%.

[13]Unfortunately, many taxi transactions happen in non-residential or tourist areas of the city, such as Midtown or Wall Street. For those census tracts, the median income is unlikely to be a good proxy for customer income.

Figure 3: Geospatial Distribution of Fares and Tip Percentages, by Pick-Up Location



Notes: Fare and Tip Percentage, by census tract of Pick-Up Location. Graphs on the left display all of NYC. Graphs on the right display Manhattan. Sample limited to Vendor-equipped cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

Table 1: Cross-Tabulation of Pick-up and Drop-off NY Borough

	Brooklyn	Manhattan	Queens	Staten Island	The Bronx	Non-NYC/Missing	Total
Brooklyn	54,927	131,159	14,496	76	220	246	201,124
Manhattan	183,845	12,793,937	188,085	170	26,055	9,976	13,202,068
Queens	58,108	170,306	139,928	88	1,226	3,472	373,128
Staten Island	70	131	25	828	2	69	1,125
The Bronx	51	2,977	397	2	8,087	173	11,687
Non-NYC/Missing	401	6,440	568	77	160	23,957	31,603
Total	297,402	13,104,950	343,499	1,241	35,750	37,893	13,820,735

Notes: Sample limited to cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

Table 2: Summary Statistics by Ride

	Competitor	Vendor	Total
Fare	9.690	9.813	9.755
	(5.480)	(5.669)	(5.580)
Tip Amount	1.694	1.920	1.813
	(1.253)	(1.431)	(1.354)
Tip as Percentage of Fare	18.27	21.46	19.95
	(9.254)	(16.59)	(13.72)
Tip Corresponds to a Default Tip Option	0.556	0.495	0.524
	(0.497)	(0.500)	(0.499)
Ride Duration (Minutes; Dropoff Time - Pickup Time)	12.67	12.85	12.77
	(8.136)	(7.931)	(8.029)
Ride Distance (Miles)	2.531	2.596	2.565
	(2.310)	(2.401)	(2.359)
Zero Tip	0.0196	0.0289	0.0245
	(0.139)	(0.167)	(0.155)
High Choice [Pr(Select 'High' Default Tip | Selected a Default Option)]	0.129	0.0371	0.0832
	(0.335)	(0.189)	(0.276)
Med Choice [Pr(Select 'Med' Default Tip | Selected a Default Option)]	0.419	0.184	0.302
	(0.493)	(0.388)	(0.459)
Low Choice [Pr(Select 'Low' Default Tip | Selected a Default Option)]	0.452	0.778	0.615
	(0.498)	(0.415)	(0.487)
Observations	6,542,783	7,277,952	13,820,735

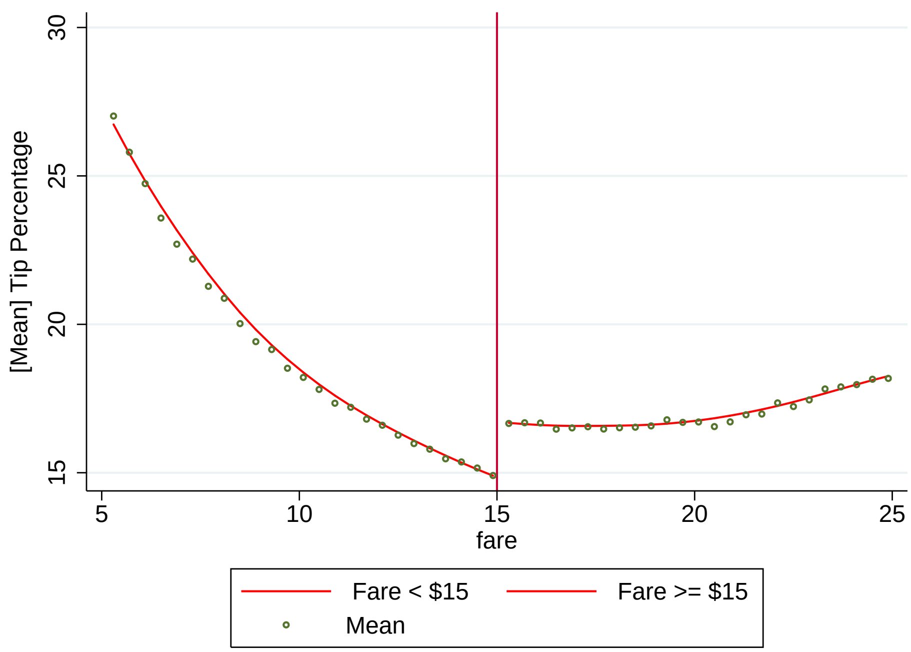
Notes: High Choice (Vendor: $4 or 30%, Competitor: 25%), Medium Choice (V: $3 or 25%, C: 20%), Low Choice (V: $2 or 20%, C: 15%) estimates are conditional on using a default tip option. Sample limited to cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

3 Regression Discontinuity

3.1 Visual Evidence

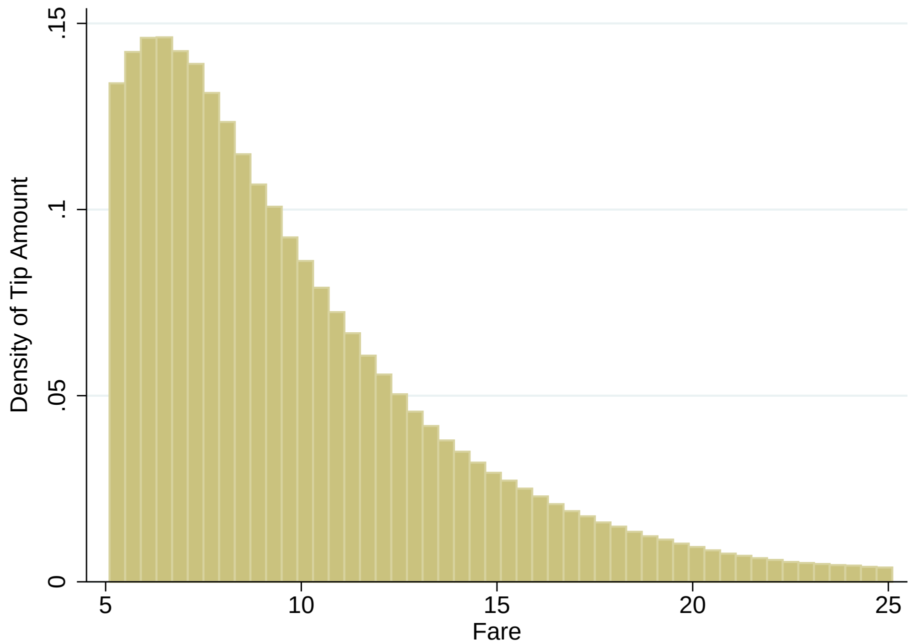
We start by presenting visual evidence of the discontinuity. We limit the forcing variable (fare) to be between \$5 and \$25 and calculate the mean tip percentage within each of the discrete fare amounts (\$0.40 increments). On each side of the discontinuity, we scatter plot these estimates and perform a lowess smoother separately on either side of the discontinuity. Figure 4 displays this plot for tip percentages on Vendor equipped cabs during pricing regime one, clearly demonstrating a jump at \$15. As a first test of the validity of the RDD, Figure 5 demonstrates that the density of the forcing variable is smooth.

Figure 4: Lowess Smoothed Mean Tip Percentages Within Each Discrete Fare Amount (\$0.40 Intervals)



Notes: Sample limited to Vendor-equipped cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

Figure 5: Histogram of Fares



Notes:Sample limited to Vendor-equipped cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

3.2 Regression

To supplement the visual evidence, we estimate a regression discontinuity model for the case of a forcing variable with discrete support. To the extent that the tip paid is smoothly related to the fare, observations at either side of the cutoff can be used to identify the causal effect of a change in the suggested amount. Following Lee and Card (2008), we estimate equation (1):

$$Y_i = \alpha \mathbb{I}\left(A_r \geq 15\right) + \beta_1 p\left(A_r - 15\right) + \beta_2 \mathbb{I}\left(A_r \geq 15\right) * p\left(A_r - 15\right) + \theta + u_r \tag{1}$$

Where Y_i is the tip amount in dollars, $\mathbb{I}\left(A_r \geq 15\right)$ is an indicator function that the fare is greater than or equal to \$15, $p\left(A_r - 15\right)$ is a polynomial in the fare centered at the discontinuity, and θ is a vector of fixed effects. We use pick-up hour, day of week, pick-up borough, and drop-off borough fixed effects. Since our source of variation is at the fare value relative to the discontinuity, rather than the ride level, we follow Lee and Card (2008) and cluster our standard errors at the level of the forcing variable, thereby correcting our degrees of freedom and allowing for random specification errors due to the discrete bins. We estimate four specifications in Table 3, starting with a 2$^{\text{nd}}$ order polynomial in the first column up to a 5$^{\text{th}}$ order polynomial in the last column. Our local treatment effect is a \$0.27 to \$0.30 increase in tip amounts over a baseline level at the cut-off of \$2.47.

Table 3: Regression Discontinuity Estimates of the Effect on Tip Amount

	(1)	(2)	(3)	(4)
	Tip Amt	Tip Amt	Tip Amt	Tip Amt
	b/se	b/se	b/se	b/se
1_[Fare>=15]	0.292***	0.276***	0.275***	0.296***
	(0.004)	(0.006)	(0.008)	(0.010)
Constant	2.392***	2.398***	2.393***	2.390***
	(0.006)	(0.006)	(0.007)	(0.007)
N	6,218,194	6,218,194	6,218,194	6,218,194
r2	0.208	0.208	0.208	0.208

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Column (1) - (4) present 2$^{\text{nd}}$-5$^{\text{th}}$ order polynomials. Robust standard errors clustered at the fare amount relative to the discontinuity (\$15). All specifications include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than \$5 and less than \$25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

3.3 Other Outcomes of Interest

Our primary outcome of interest, tip amount, is produced through movements along an extensive margin (using a default suggestion) and two intensive margins (amounts tipped either manually or through one of the suggestions). Table 4 presents results for a number of other variables.

Table 4: Regression Discontinuity Estimates of the Effect on Alternative Outcomes

	(1) Tip Percent b/se	(2) Default Tip b/se	(3) Default Tip Amt b/se	(4) Manual Tip Amt b/se	(5) High Choice b/se	(6) Med Choice b/se	(7) Low Choice b/se
1_[Fare>=15]	2.024***	-0.243***	0.714***	0.368***	-0.021***	-0.169***	0.190***
	(0.038)	(0.002)	(0.003)	(0.011)	(0.001)	(0.002)	(0.003)
Constant	15.574***	0.783***	2.596***	1.524***	0.100***	0.400***	0.500***
	(0.056)	(0.003)	(0.004)	(0.013)	(0.002)	(0.003)	(0.003)
N	6,218,194	6,218,194	3,227,733	2,990,461	3,227,733	3,227,733	3,227,733
r2	0.098	0.058	0.568	0.122	0.015	0.029	0.038

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

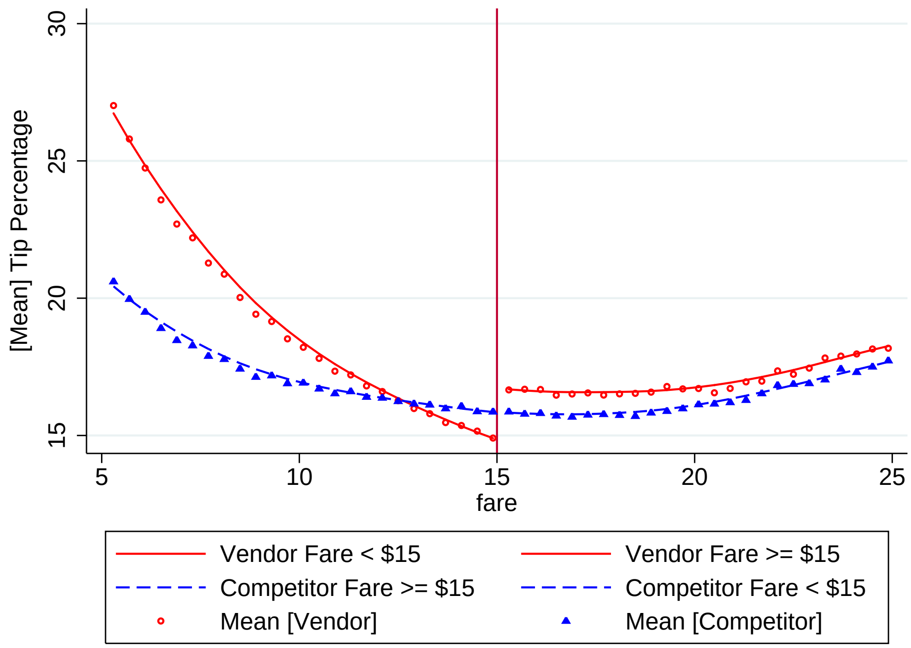
Notes: Columns 1 (Tip Percent), 3 (Default Tip Amount), and 4 (Manual Tip Amount) use continuous outcome variables, while columns 2 (Default Tip), 5 (High Choice), 6 (Medium Choice), and 7 (Low Choice) use binary outcome variables. High Choice (Vendor: $4 or 30%, Competitor: 25%), Medium Choice (V: $3 or 25%, C: 20%), Low Choice (V: $2 or 20%, C: 15%) estimates are conditional on using a default tip option. Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3$^{\text{rd}}$-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Notably, the higher tip suggestions induce a 24 percentage point reduction in the probability of using a default suggestion, and a shift in the composition of those that use default suggestions toward the low option. However, since the low option is approximately equal to the medium option to the left of the discontinuity, we see a net increase in the amount tipped by those that select a default option. There is also an increase in amount tipped manually, which reflects a change in the composition of the tippers, but may also reflect the influence of the higher suggestions on those that would tip manually when faced with either set of suggestions.

3.4 Robustness

As our first robustness test, we repeat the visual RD graph for the Competitor. Figure 6 shows that the jump is absent for the Competitor. We re-estimate regression specification (1), changing the outcome variable to ones that should not be significantly affected by the default suggestions. Table 5 shows that the treatment effects are small and in conflicting directions for ride distance and duration, and the effects on passenger count, hour of the day, and day of the week are insignificant.

Figure 6: Lowess Smoothed Mean Tip Percentages Within Each Discrete Fare Amount ($0.40 Intervals), for Vendor and Competitor



Notes: Sample limited to cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

Table 5: Falsification Test 1: Regression discontinuity for trip distance, ride duration, hour of pick-up, day of the week, and passenger count

	(1) Distance b/se	(2) Duration b/se	(3) Hour of Pick-Up b/se	(4) Day of the Week b/se	(5) Passenger Count b/se
1_[Fare>=15]	-0.015***	0.147***	-0.011	0.009	0.003
	(0.005)	(0.024)	(0.012)	(0.009)	(0.002)
Constant	5.661***	15.409***	12.836***	3.096***	1.968***
	(0.004)	(0.023)	(0.014)	(0.011)	(0.003)
N	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194
r2	0.884	0.780	0.378	0.059	0.901

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3rd-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough (except when one of these is the outcome variable]. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Table 6: Falsification Test 2: Regression Discontinuity for Placebo Integers.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt	Tip Amt
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
1_[Fare>=9]	0.058^{***}												
	(0.003)												
1_[Fare>=10]		0.030^{***}											
		(0.003)											
1_[Fare>=11]			0.007^{**}										
			(0.003)										
1_[Fare>=12]				-0.041^{***}									
				(0.003)									
1_[Fare>=13]					-0.120^{***}								
					(0.004)								
1_[Fare>=14]						-0.056^{***}							
						(0.004)							
1_[Fare>=15]							0.276^{***}						
							(0.006)						
1_[Fare>=16]								0.092^{***}					
								(0.006)					
1_[Fare>=17]									-0.030^{***}				
									(0.009)				
1_[Fare>=18]										-0.111^{***}			
										(0.009)			
1_[Fare>=19]											-0.064^{***}		
											(0.015)		
1_[Fare>=20]												-0.184^{***}	
												(0.014)	
1_[Fare>=21]													-0.145^{***}
													(0.026)
Constant	1.912^{***}	1.994^{***}	2.076^{***}	2.166^{***}	2.247^{***}	2.319^{***}	2.398^{***}	2.724^{***}	2.999^{***}	3.237^{***}	3.417^{***}	3.643^{***}	3.798^{***}
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
N	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194	6,218,194
r2	0.207	0.207	0.207	0.207	0.207	0.207	0.208	0.207	0.207	0.207	0.207	0.207	0.206

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

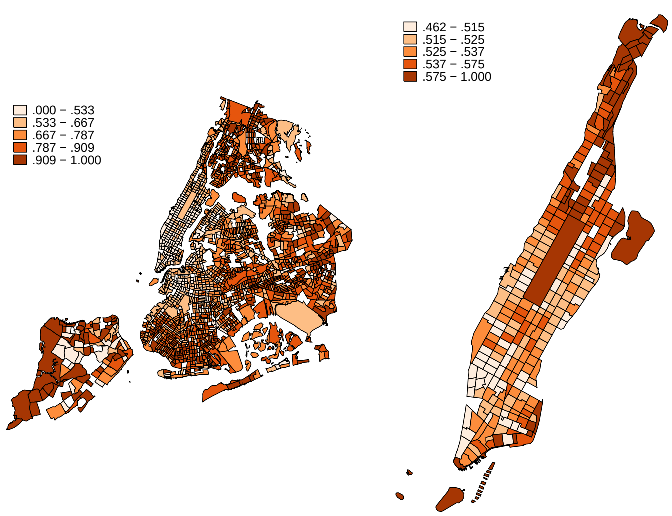
Notes: Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3$^{\text{rd}}$-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Finally, we estimate the RDD model at placebo integers, using a 3rd order polynomial in all specifications. We see in Table 6 that while most of these values are precisely estimated, they are smaller in magnitude than the $15 discontinuity, and often negative.

4 Comparing Across Vendors

While our regression discontinuity design provides compelling identification, it is limited to a localized treatment effect. Furthermore, interpretation is complicated by the possibility that customers may have difficulty translating the dollar amounts into percentages, and vice versa. One way to expand upon our results would be to compare rides over which both credit card machine companies provided percentage default suggestions. For fares above $15, the Vendor provided suggestions of 20%, 25%, and 30%, while the competitor provided suggestions of 15%, 20%, and 25%. However, Figure 7 shows that the geospatial distribution of pick-up locations differs between the two companies, and that tip percentages also vary by pick-up location.

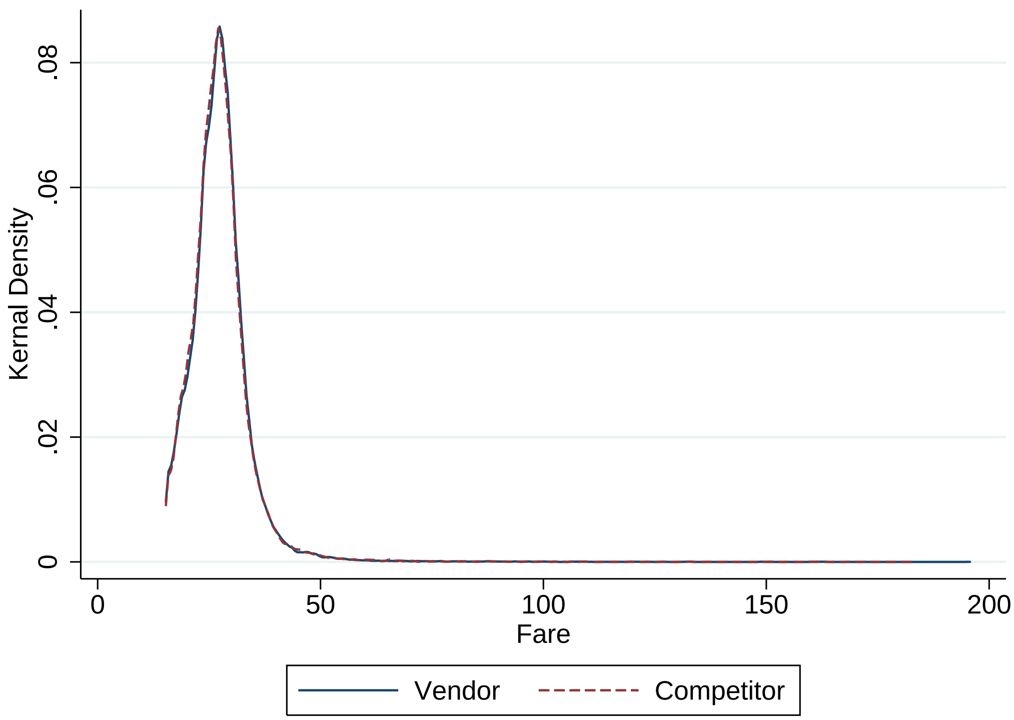
Figure 7: Proportion of Rides Originating with a Vendor versus a Competitor Equipped Cab, By census tract of Pick-Up Location



Notes: Graphs on the left display all of NYC. Graphs on the right display Manhattan. Sample limited to Vendor-equipped cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

While we can control for the pick-up and drop-off location, there may be other unobservable differences in driver-customer match that affect the tip amounts. To address this challenge, we limit our analysis to rides that originate at LaGuardia airport.[14] Customers queue at lines that contain a mix of taxis equipped with both credit card machine companies. Figure 8 shows that the distribution of fares is comparable across the two credit card companies when we limit the sample to rides that are above \$15 and originate at LaGuardia.[15]

Figure 8: Distribution of fares across the Vendor and Competitor Equipped Taxis



Notes: The sample is limited to rides under the first pricing regime (January 1 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday/Sunday), for fares greater than or equal to \$15, and only those rides that started at the census tract associated with LaGuardia Airport (331).

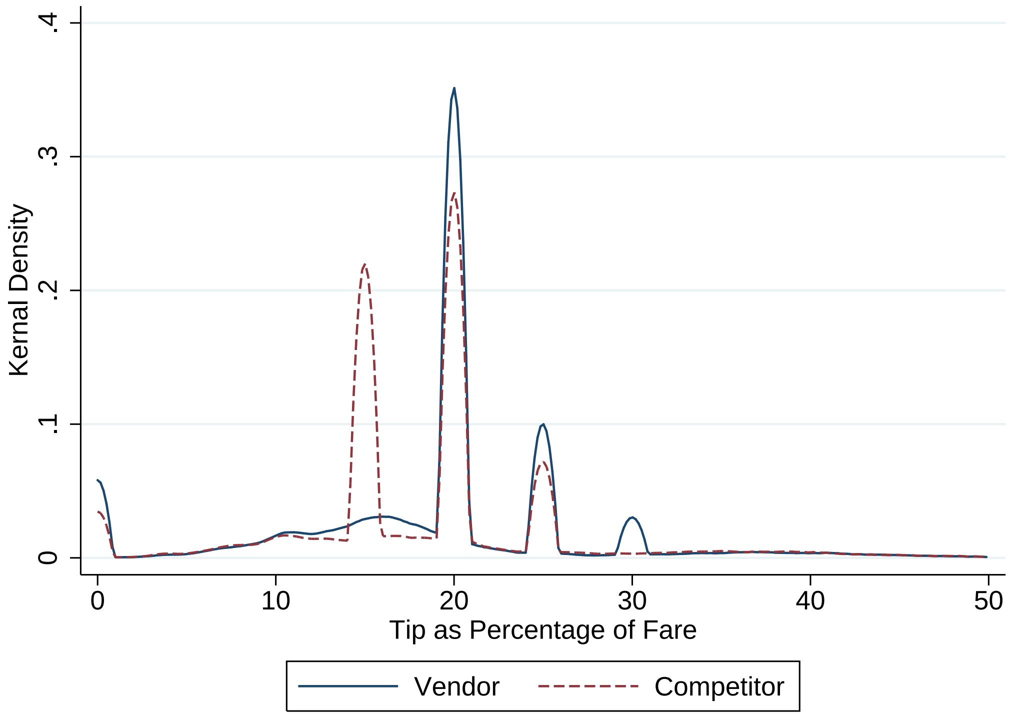
For the distribution of tip percentages, we limit the sample to fares with tip percentages less than or equal to 50% in order to provide greater visual clarity. Figure 9 demonstrates

[14]We exclude JFK airport because the majority of rides use a \$45 fixed fare, complicating our placebo test of equality in fares between vendors. In the web appendix, we repeat the analysis in this section pooling both LGA and JFK observations. Our estimates in that pooled sample are qualitatively similar.

[15]However, it should be noted that a simple t-test of fare across the vendor (27.47) and competitor (27.36) rejects equality ($p = .0089$).

the stark difference in the two distributions, with the higher set of defaults inducing a distribution that has significantly more density around its lowest option. The left tail of the figure is also larger for the Vendor; however, this effect is limited to zero-valued tips.

Figure 9: Distribution of Tip Percentages Across the Vendor and Competitor Equipped Taxis



Notes: The sample is limited to rides under the first pricing regime (January 1 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday/Sunday), for fares greater than or equal to $15, and only those rides that started at the census tract associated with LaGuardia Airport (331).

We provide a regression analysis of these effects in Table 7. To address concerns of any type of sorting between drivers and credit card machine companies, we also provide specifications with fixed effects for driver, pick-up hour, and borough of the drop-off location. These restricted specifications isolate our estimates to the 19% of drivers in our sample that drove on cars equipped by both of the vendors, allowing us to identify the coefficient on "vendor" while controlling for time-invariant driver characteristics. We find that the increase in fare is significant at the 5% level in specifications that do not include fixed effects

(Column 1); however, this difference is insignificant in the specification that controls for time-invariant driver characteristics (Column 2). We find a significant increase in the tip percentage, consistent with our regression discontinuity estimates, though the effect size is smaller in magnitude. We also find a reduction in the probability of using one of the suggested amounts, also consistent with Section 3.

Table 7: OLS - Comparison of Vendor (20%/25%/30%) and Competitor (15%/20%/25%) - Fare, Tip Percentage, and Default Tip

	(1) Fare b/se	(2) Fare b/se	(3) Tip Percent b/se	(4) Tip Percent b/se	(5) Default Tip b/se	(6) Default Tip b/se
Vendor	0.117*	0.171	0.646***	0.713***	-0.084***	-0.078***
	(0.062)	(0.147)	(0.099)	(0.259)	(0.004)	(0.013)
Constant	27.358***	28.704***	18.700***	18.867***	0.632***	0.613***
	(0.049)	(0.750)	(0.076)	(1.454)	(0.003)	(0.078)
Fixed Effects		X		X		X
N	101,710	18,184	101,710	18,184	101,710	18,184
r2	0.000	0.292	0.001	0.277	0.007	0.229

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the driver level. Even columns include fixed effects for driver, pick-up hour, and drop-off borough. Columns 5 & 6 are linear probability models.

5 Cost of High Defaults

Our results in sections 3 & 4 highlight the revenue benefits of setting higher defaults. In those two sections, we also show that high defaults result in a higher proportion of customers foregoing the suggested amounts and instead keying in their own manual tip amount. However, we find an increase in the manual tip average, and so this channel alone may not be sufficient to reduce tips. In this section, we document an unambiguous cost of higher defaults: the potential to discourage tipping altogether.

We first consider the across-vendor analysis of Section 4. Figure 9 documents a larger density around the zero tip percentage for Vendor equipped taxi cabs. Table 8 repeats the regression analysis of Table 7 with a binary outcome variable for leaving no credit card tip and one for leaving a small tip. Columns 3 and 4 show a small decrease in the probability of leaving a tip greater than 0% but less than 10%; however, columns 1 and 2 show a significant increase in the probability of leaving a zero-valued tip. This increase in the proportion of zero-valued tips has some parallel in the vast literature on ultimatum games. Insofar as customers have some fixed notion of a "fair" tip, presenting the higher suggestions might have led them to punish the drivers by leaving a lower tip than would be provided in the absence of the "unfair" split. This result highlights a potential cost of setting defaults too high, although we cannot confirm whether this cost would persist in a context featuring homogeneous suggestions across vendors.[16] The backlash to high suggestions may hinge on the existence of a reference "fair tip" in a comparable market. Furthermore, without making strong assumptions, we cannot use this cost to trace out the set of optimal default suggestions.

Table 8: OLS - Comparison of Vendor (20%/25%/30%) and Competitor (15%/20%/25%)- Zero-Valued Tip

	(1) Zero Tip b/se	(2) Zero Tip b/se	(3) TipPercent0to10 b/se	(4) TipPercent0to10 b/se
Vendor	0.028***	0.028***	-0.003**	-0.016***
	(0.001)	(0.006)	(0.001)	(0.006)
Constant	0.039***	0.027	0.052***	0.105**
	(0.001)	(0.026)	(0.001)	(0.044)
Fixed Effects		X		X
N	101,710	18,184	101,710	18,184
r2	0.004	0.220	0.000	0.214

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the driver level. Even columns include fixed effects for driver, pick-up hour, and drop-off borough. The dependent variable in columns 1 & 2 is an indicator for whether the customer left a zero-valued tip, while the dependent variable in columns 3 & 4 is an indicator for leaving a tip that is greater than 0% and less than 10% of the fare.

There are a few possible alternative explanations of the increase in zero-valued tips. First,

[16] For example, all vendors currently offer the 20%/25%/30% suggestions.

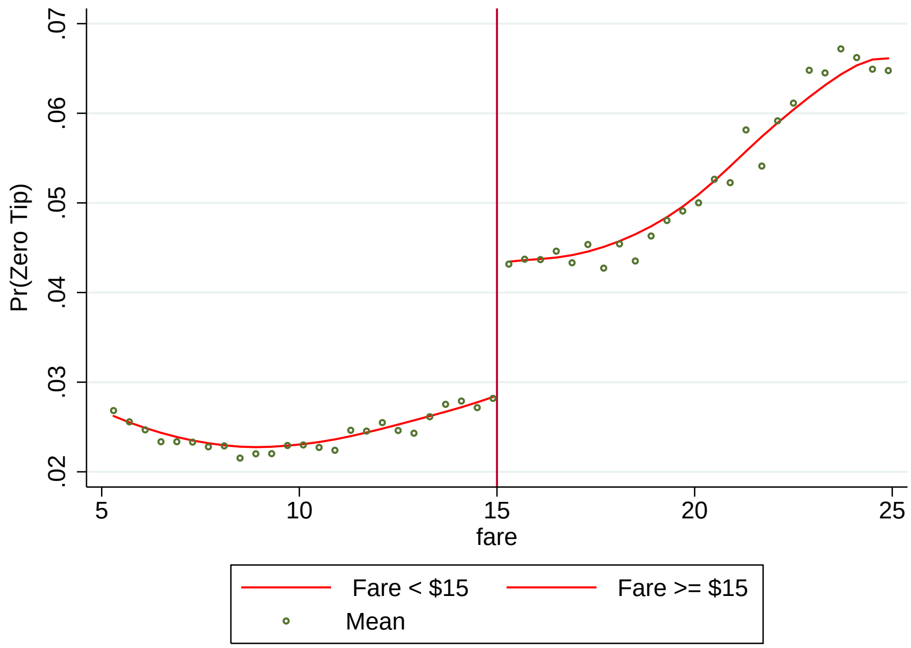
presenting customers with the 20%/25%/30% buttons (and the ability to key in any amount), rather than the 15%/20%/25% buttons (and again the ability to key in any amount), may have caused them to leave their tips in cash, for which we do not have data, rather than credit (despite still opting to pay the fare in credit). A second, more plausible alternative hypothesis is that these zero-valued tip entries reflect data errors that were more likely to be produced by Vendor machines. In cleaning the data, we removed all zero-valued distance and ride duration observations; however, we found that there were slightly more of these distance errors associated with the Competitor (0.88% vs. 0.65%) and more ride duration errors associated with the Vendor (.65% vs. .05%). It is possible that some of the zero-valued tip percentage entries are residual electronic errors or tests, and that these tip errors are more concentrated in Vendor credit card machines. Finally, a third hypothesis is that our result is driven by possible differences in how the suggestions were presented by the Competitor and Vendor.[17]

To rule out these last two alternative hypotheses, we again return to the regression discontinuity specification. Figure 10 shows that this behavioral response is preserved when looking within Vendor. Our regression estimate of the effect is a significant 1.7 percentage point increase in the probability of leaving a zero-valued tip.[18] These results suggest that we are detecting a true negative behavioral response to the higher suggestions.

[17]Currently, the Vendor provides a visible conversion of the percentage into dollar amounts, while the Competitor does not present this information. We do not know if this difference in information display was present in 2009.

[18]This specification is similar to those used in Table 4 (i.e. a 3rd degree polynomial with fixed effects for driver, day, hour, and boroughs of the pick-up and drop-off locations). The constant in the regression is 0.039.

Figure 10: Lowess Smoothed Mean of "Zero-Valued Tip" Within Each Discrete Fare Amount ($0.40 Intervals)



Notes: Sample limited to cab rides during the first price regime: January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday.

6 Default Mechanisms and Heterogeneity

We consider three complementary explanations for the observed default effect. First, customers may be rationally inattentive if the cognitive effort or time costs are sufficiently high to justify the additional tip. A second possible explanation is that the default serves an information transmission purpose, acting as a signal of the social norm to unfamiliar customers.[19] Finally, customers may experience disutility from deviating from the status quo,

[19]Similarly, it is also possible that the tip suggestions themselves change the prevailing social norm, even among experienced customers. For instance, on November 2008 the NY Times reported that "*tips, which hovered around 10 percent when cab rides were cash only, averaged 22 percent on credit-card transactions this fall.*" (Source: http://www.nytimes.com/2009/11/08/nyregion/08taxi.html) Given our lack of cash tip data (and our limitation to 2009 data), we cannot substantiate this claim, nor can we comment on the long-term effects of a change in the prevailing suggestion; however, a shift in norms is one plausible hypothesis for this informal finding. The long-term effect of default policy changes is an interesting subject that merits further research.

either due to social pressure or other forms of psychological resistance.[20] Our data do not allow us to cleanly parse the three proposed mechanisms; however, we present a few tests that attempt to shed light on their relative roles.

We first consider the information transmission mechanism. To rule out that the possibility that the default effect is entirely driven by information transmission, we split the sample into Manhattan and other Boroughs. The purpose of this split is to reduce the concentration of tourists, limiting us to riders that are more experienced with taxi cabs and the specific NYC taxi tipping norms. An important caveat, which will apply to all of our heterogeneity tests, is that we may not be isolating the proposed mechanism. Both the types of rides that originate outside of Manhattan and the types of riders who live outside of Manhattan are different along a number of possible dimensions (e.g. different social norms), and thus we will potentially be conflating differences in information with other observable and unobservable offsetting differences. With this critical caveat in mind, we proceed to estimate these heterogeneous treatment effects. Table 9 shows that the default effect is of similar magnitude across rides that originate inside Manhattan (Column 1) and rides that originate outside of Manhattan (Column 2).

[20]A prime example of this social pressure mechanism is expressed in a New York Sun magazine article on the introduction of the credit card system: *"It forces you to tip," a Manhattan resident who recently tipped 15% on a $14 fare, Greg Mack, said. "What if you didn't enjoy the ride? It made me feel obligated."*(Source: http://www.nysun.com/business/hot-tip-for-cabbies-credit-cards-boost-tips/72783/).

Table 9: Heterogeneity by Pick-up/Drop-off Location: Regression Discontinuity Estimates of Default Effect on Tip Amount by Manhattan vs. Other Boroughs

	(1) Manhattan b/se	(2) NonManhattan b/se
1_[Fare>=15]	0.271***	0.272***
	(0.006)	(0.026)
Constant	2.324***	2.293***
	(0.006)	(0.024)
N	5,987,793	236,007
r2	0.194	0.323

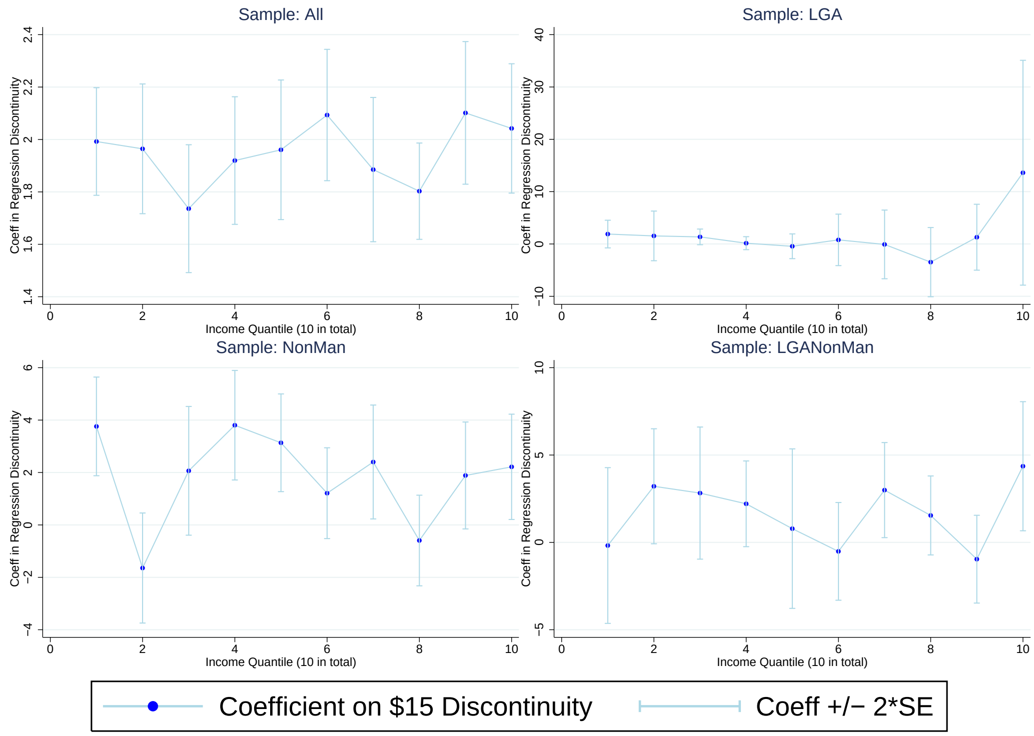
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3rd-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday). Column 1 is limited to rides that originate inside Manhattan, while Column 2 is limited to rides that originate outside Manhattan.

We next turn to the rational inattention explanation. Our first exercise examines how the default effect varies with income. Because wealthier riders have a lower marginal utility of income, they have less incentive to be attentive to the shift. Similarly, these customers will potentially have higher time costs. This reasoning suggests that higher income customers should (rationally) exhibit a larger default effect. Alternatively, wealthier individuals may have greater access to distraction-reducing devices, allowing them to deplete less attention during the day (Banerjee and Mullainathan 2008), and thus be less susceptible to default effects. A wealthier customer may also be more likely to take more taxi rides, and default effects possibly attenuate with experience (Lofgren et al 2012). Finally, income may be unrelated to the default effect either due to a negligible time cost of the task or due to the possible orthogonality of cognitive effort costs and income. However, using the underlying income may not necessarily isolate differences in time costs or the marginal utility of income. For example, drivers may expect lower tips from riders in lower income areas, and thus the riders may experience less disutility in deviating from the default suggestion. Ultimately, this exercise is only weakly suggestive of the rational inattention mechanism; however, it may be an intrinsically interesting source of heterogeneity. For example, Goldin and Homonoff (2012) found that low-income customers were more attentive to a low salience cigarette tax than were high income customers. In contrast, Beshears et al (2010) found that 401(k) savings defaults had a greater influence on low-income employees than on high-income employees.

To proxy income, we use a variety of different sample groups. For the full sample, we proxy customer income by the average of the median income associated with the pick up location and drop-off location census tracts. These pick-up and drop-off locations would not be an accurate assessment of tourists or any other customers that do not start or end at their home address. To partially reduce this concern, we use a variety of other specifications that attempt to remove these non-representative customers. One set of specifications limit the sample to rides that both start and end outside Manhattan. Another set of specifications limit to rides that either start or end at LaGuardia airport (proxying income by the median income in the pick-up location census tract if the ride ends at LGA or by the income at the drop-off location if the ride starts at LGA). Finally, the most conservative set of specifications limits to rides that start at LaGuardia airport and end outside Manhattan. We then split these rides into ten income quantiles and run the regression discontinuity for each of these sub-samples. Figure 10 plots the coefficients from these regressions, finding no systematic pattern in income. Although it might be that the absence of a clear pattern is due to measurement error in the income variable, the constancy in the discontinuity suggests that default effects are similar for customers across the different proxied income brackets in our sample.

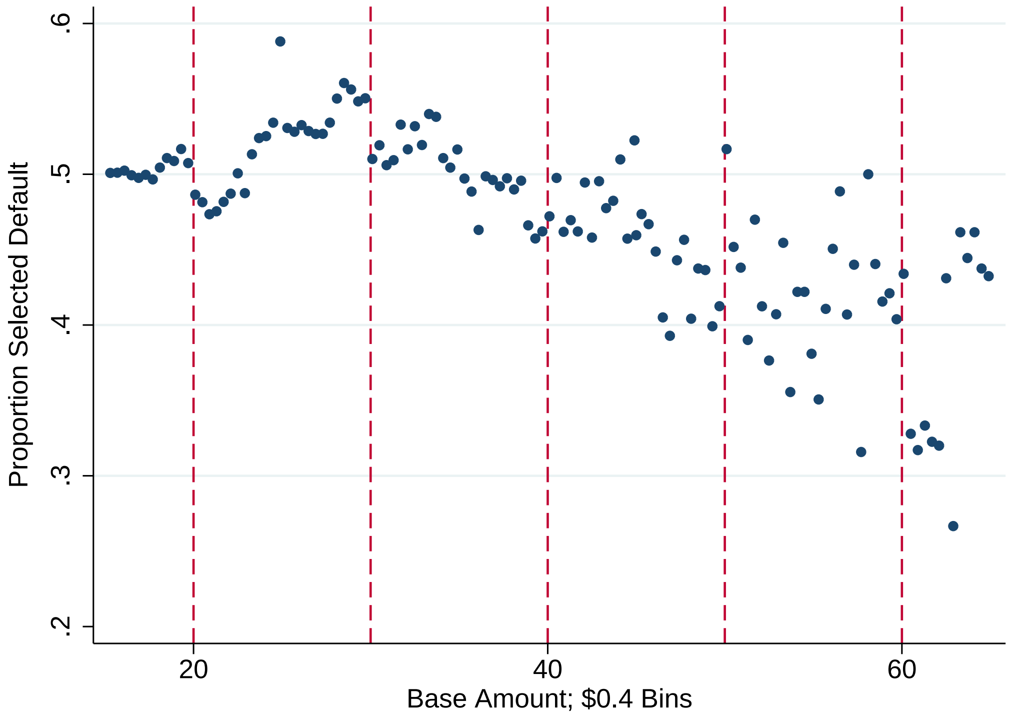
Figure 11: Coefficients from Regression Discontinuity Estimates on Income Quantile Sub-Samples



Notes: Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3rd-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Finally, to investigate the hypothesis that cognitive effort costs drive reliance on the default suggestions, we plot the usage of a default tip button at different fare. We hypothesize that usage of a default button should decrease around numbers upon which it is easier to perform mental computations, e.g. $20. Figure 11 does not provide clear evidence consistent with this hypothesis.

Figure 12: Computational Costs: Proportion of Customers that Use a Default Option by Each \$0.40 Fare Bin



Notes: The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than \$15 and lower than \$65 and during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

7 Conclusion

Using an extensive dataset, we show that a small change in default tip suggestions has a significant effect on tipping amounts. Our data allow us to provide very clean identification in a large naturalistic field setting. We use a regression discontinuity design to show that an upward shift in the set of suggestions induces higher average tip amounts, despite significantly reducing the probability that customers use one of the defaults. To address concerns about differences in the framing of suggestions (dollar amounts vs. percentages), as well as to provide less localized treatment effect estimates, we performed a secondary analysis of trips originating at the airport. Exploiting these quasi-random driver-to-customer matches, we again find that higher default tip suggestions (20%/25%/30% vs. 15%/20%/25%) result in higher average tip amounts. This analysis also reveals a potential cost of setting defaults too high – customers are also more likely to leave no tip in response to the higher defaults. Finally, we highlight potential mechanisms and show a surprising dearth of heterogeneity in our treatment effects. As firms begin to use the insights of behavioral economics to inform their product design and promotion, our study suggests that default effects can be exploited even in habitually encountered consumer choices. However, there may be a backlash to defaults that exceed certain thresholds, and so firms and policy makers alike should be cognizant of this potential cost when optimally designing their defaults.

References

[1] Alberto Abadie and Sebastien Gay. The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-Country Study. *Journal of Health Economics*, 25(4):599–620, July 2006.

[2] Ofer Azar. Business Strategy and the Social Norm of Tipping. *Journal of Economic Psychology*, 32(3):515–525, June 2011.

[3] Abhijit V Banerjee and Sendhil Mullainathan. Limited Attention and Income Distribution. *American Economic Review*, 98(2):489–493, April 2008.

[4] John Beshears, James Choi, David Laibson, and Brigitte Madrian. The Limitations of Defaults. *Working Paper*, 2010.

[5] Erin Bronchetti, Thomas Dee, and David Huffman. When a Nudge Isn't Enough: Defaults and Saving Among Low-Income Tax Filers. *NBER Working Paper*, 2011.

[6] Gabriel Carroll, James Choi, David Laibson, Brigitte Madrian, and Andrew Metrick. Optimal Defaults and Active Decisions. *Quarterly Journal of Economics*, 124(4):1639–1674, November 2009.

[7] James Choi, David Laibson, and Brigitte Madrian. Plan Design and 401 (k) Savings Outcomes. *National Tax Journal*, LVII(2):275–298, 2004.

[8] Henrik Cronqvist and Richard Thaler. Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience. *American Economic Review*, 2004.

[9] Stefano DellaVigna and Ulrike Malmendier. Paying Not to Go to the Gym. *American Economic Review*, 2006.

[10] Sibilla Di Guida, Davide Marchiori, and Ido Erev. Decisions Among Defaults and the Effect of the Option to Do Nothing. *Economics Letters*, 117(3):790–793, December 2012.

[11] Julie Downs, George Loewenstein, and Jessica Wisdom. Strategies for Promoting Healthier Food Choices. *American Economic Review*, 99(2):159–164, April 2009.

[12] Jacob Goldin and Tatiana Homonoff. Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity. *American Economic Journal: Economic Policy [Forthcoming]*, pages 1–59, 2012.

[13] Eric Johnson, Steven Bellman, and Gerald Lohse. Defaults, Framing and Privacy: Why Opting In-Opting Out. *Marketing Letters*, 2002.

[14] Eric Johnson and Daniel Goldstein. Do Defaults Save Lives? *Science*, pages 4–5, 2003.

[15] Eric Johnson, John Hershey, Jacqueline Meszaros, and Howard Kunreuther. Framing, Probability Distortions, and Insurance Decisions. *Journal of Risk and Uncertainty*, 1993.

[16] Daniel Kahneman. *Thinking, Fast and Slow*, volume 1. Farrar, Straus and Giroux, 2011.

[17] Dean Karlan, John List, and Eldar Shafir. Small Matches and Charitable Giving: Evidence from a Natural Field Experiment. *Journal of Public Economics*, 2011.

[18] Punam Anand Keller, Bari Harlam, George Loewenstein, and Kevin G. Volpp. Enhanced Active Choice: A New Method to Motivate Behavior Change. *Journal of Consumer Psychology*, 21(4):376–383, October 2011.

[19] David S. Lee and David Card. Regression Discontinuity Inference with Specification Error. *Journal of Econometrics*, 142(2):655–674, February 2008.

[20] Asa Lofgren, Peter Martinsson, Magnus Hennlock, and Thomas Sterner. Are Experienced People Affected by a Pre-det Default Option – Results from a Field Experiment. *Journal of Environmental Economics and Management*, 63(1):66–72, January 2012.

[21] Brigitte Madrian and Dennis Shea. The Power of Suggestion: Inertia in 401 (k) Participation and Savings Behavior. *Quarterly Journal of Economics*, CXVI(November), 2001.

[22] Richard Thaler and Cass Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press, New Haven, 2008.

A Appendix

Our final dataset was constructed by first performing a number of consistency checks and then removing data that appeared to be generated by electronic tests or other types of data errors. The full sample of 170,896,479 was reduced to 13,820,735 observations by performing a number of procedures. First, we made the following consistency adjustments:

1. The pick-up came after the drop-off time in 0.14% (241,965) of observations. We replaced these pick-up times with their drop-off times, and vice-versa.

2. The drop-off time came after the pick-up time of the subsequent trip in 0.36% (618,945) of observations. We set the drop-off time equal to the pick-up time of the subsequent trip for all of these cases.

The full sample of 170,896,479 was reduced to 163,348,802 by dropping all observations for which:

1. The payment type was “No Charge” (509,194; .30%) or “Dispute” (94,784; .06%).

2. The ride duration was either equal to zero or longer than 3 hours (619,468; 0.36%).

3. The distance was either equal to zero or greater than 100 miles (929,480; 0.55%).

4. The surcharge was greater than $1 (75,295; 0.04%).

5. Corresponding to drivers that drove fewer than 100 rides in 2009 (58,494; 0.03%).

6. Multiple cars were associated with the same driver during the same shift (1,298,431; 0.77%).

7. The driver’s shift was longer than 20 hours (3,872,210; 2.31%).

8. The driver’s shift was shorter than 30 minutes (89,616; 0.05%).

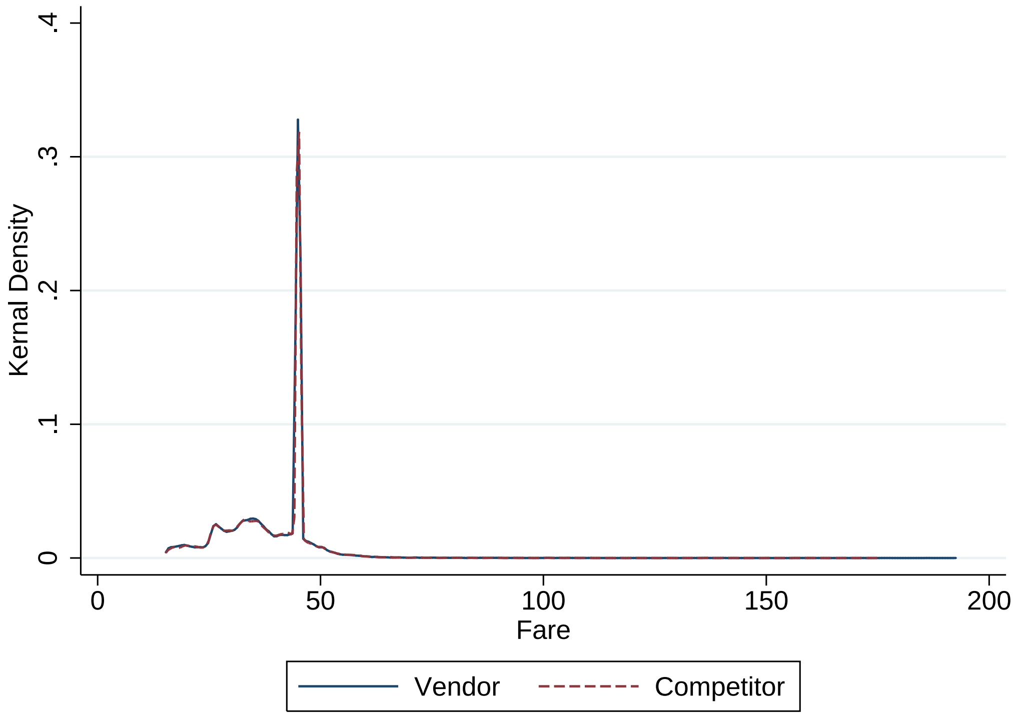
We then dropped the 5.95% (9,719,003) of observations that were on cars equipped with the third credit card machine vendor. Next, we dropped observations for which either the pick-up location or drop-off location could not be mapped to a census tract in NY, NJ, CT, or PA (2,022,206; 0.13%). To ensure that the regression discontinuity is identified off representative rides, we dropped all rides that had toll amounts applied.[21] This dropped the 4,882,697 (3.22%) rides which were associated with a toll amount greater than zero. We then made the largest sample reduction, removing the 108,620,392 rides paid by cash, as the data did not include tip information for these rides. From this sample of 38,104,504 rides, we further limited to those rides for which the base amount (the sum of the fare, tolls, surcharge, and tax) was equivalent to the "fare". Performing this reduction ensured that rides on either side of the discontinuity were comparable in terms of the time of day, time of year, and the fees faced by the customer. This reduction left 13,936,381 rides that occurred prior to November 1, 2009 and between the hours of 6am to 4pm on Monday through Friday or 6am to 8pm on Saturday and Sunday. Finally, we removed rides that didn't correspond to a multiple of \$0.40 (the unit of fare accrual) added to \$2.50 (the flat entry fee), leaving 13,820,735 observations in the final sample.

[21] Furthermore, including tolls complicates the comparison of Competitor and Vendor tip amounts, as the Competitor default tips were computed on the total base amount (including Tolls), while the Vendor default tips were computed only on the fare.

B Web Appendix

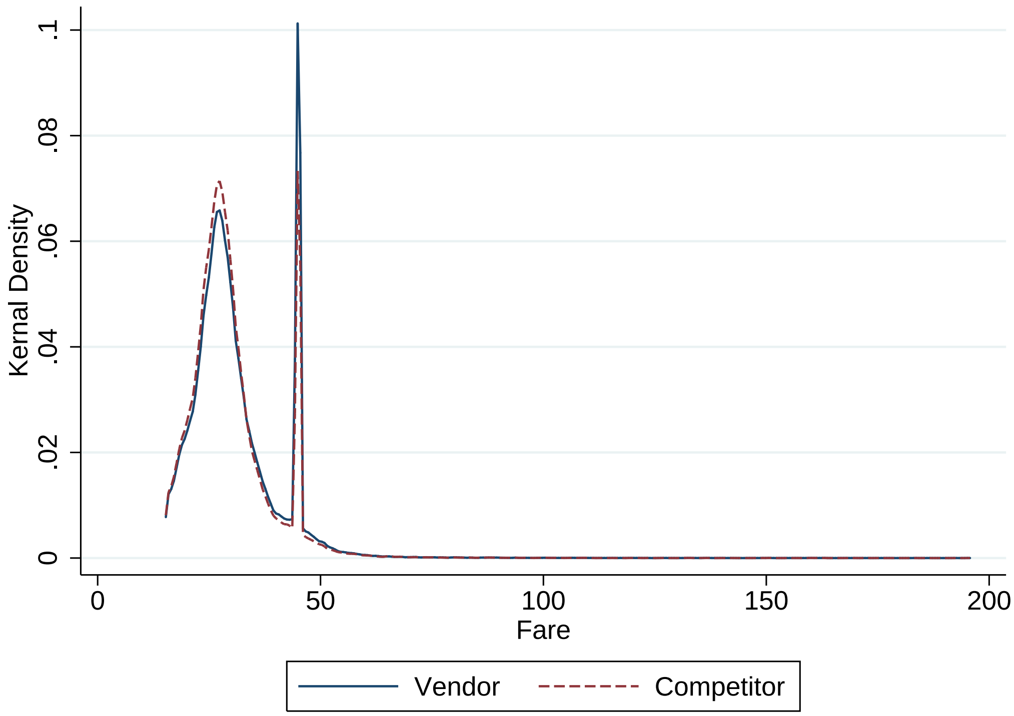
B.1 Comparing Across Vendors

Figure B.1: JFK Sample: Distribution of fares across the Vendor and Competitor equipped taxis



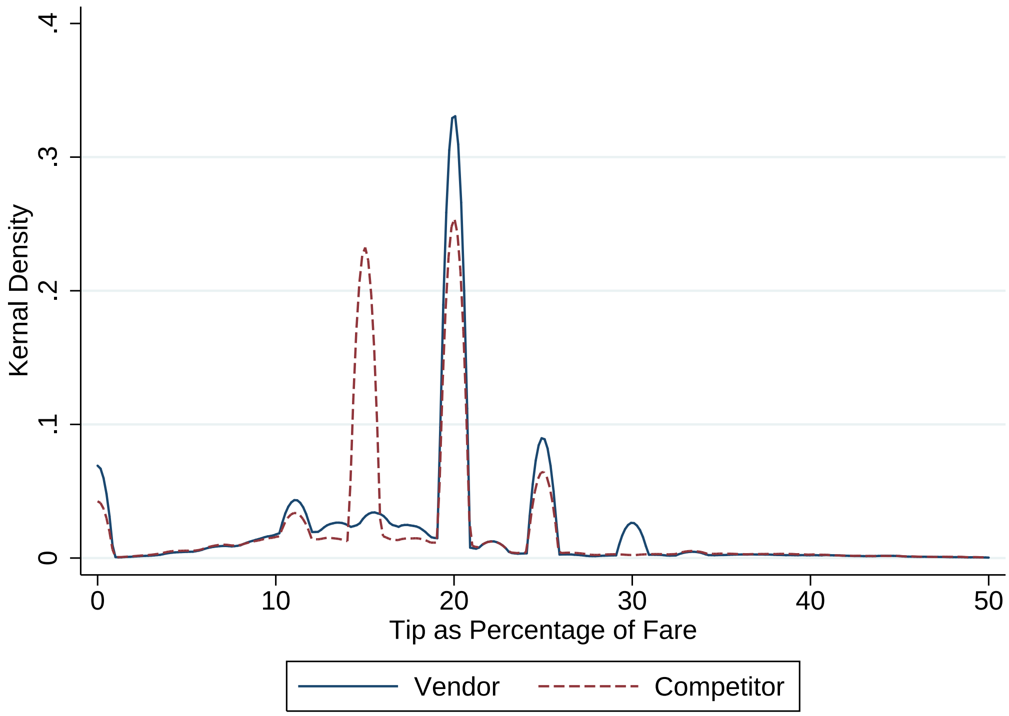
Notes: The sample is limited to rides under the first pricing regime (January 1 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday/Sunday), for fares greater than or equal to $15, and only those rides that started at the census tracts associated with JFK Airport.

Figure B.2: LGA & JFK Pooled Sample: Distribution of Fares Across the Vendor and Competitor Equipped Taxis



Notes: The sample is limited to rides under the first pricing regime (January 1 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday/Sunday), for fares greater than or equal to $15, and only those rides that started at the census tracts associated with LaGuardia Airport and JFK Airport.

Figure B.3: LGA & JFK Pooled Sample: Distribution of Tip Percentages Across the Vendor and Competitor Equipped Taxis



Notes: The sample is limited to rides under the first pricing regime (January 1 - October 31, 2009; 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday/Sunday), for fares greater than or equal to $15, and only those rides that started at the census tracts associated with LaGuardia Airport and JFK Airport.

Table B.1: LGA & JFK Pooled Sample: OLS - Comparison of Vendor (20%/25%/30%) and Competitor (15%/20%/25%) customer outcomes

	(1) Fare b/se	(2) FareFE b/se	(3) TipPercent b/se	(4) TipPercentFE b/se	(5) DefaultTip b/se	(6) DefaultTipFE b/se	(7) TipPercent0to10 b/se	(8) TipPercent0to10FE b/se	(9) ZeroTip b/se	(10) ZeroTipFE b/se
Vendor	0.732***	0.516***	0.498***	0.564***	-0.104***	-0.094***	-0.003**	-0.009**	0.032***	0.028***
	(0.109)	(0.165)	(0.075)	(0.177)	(0.003)	(0.010)	(0.001)	(0.004)	(0.001)	(0.004)
Constant	31.957***	36.606***	17.651***	16.383***	0.619***	0.639***	0.065***	0.096***	0.048***	0.055***
	(0.082)	(0.349)	(0.057)	(0.333)	(0.002)	(0.018)	(0.001)	(0.010)	(0.001)	(0.009)
N	176,220	30,006	176,220	30,006	176,220	30,006	176,220	30,006	176,220	30,006
r2	0.001	0.349	0.001	0.203	0.011	0.163	0.000	0.141	0.004	0.152

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the driver level. Even columns include fixed effects for driver, pick-up hour, and drop-off borough. Columns 5-8 are linear probability models. The dependent variable in columns 5 & 6 is an indicator for leaving a tip that is greater than 0% and less than 10% of the fare, while the dependent variable in columns 7 & 8 is an indicator for whether the customer left a zero-valued tip. The sample is limited only those rides that started at the census tracts associated with LaGuardia Airport and JFK Airport.

B.2 Default Mechanisms and Heterogeneity

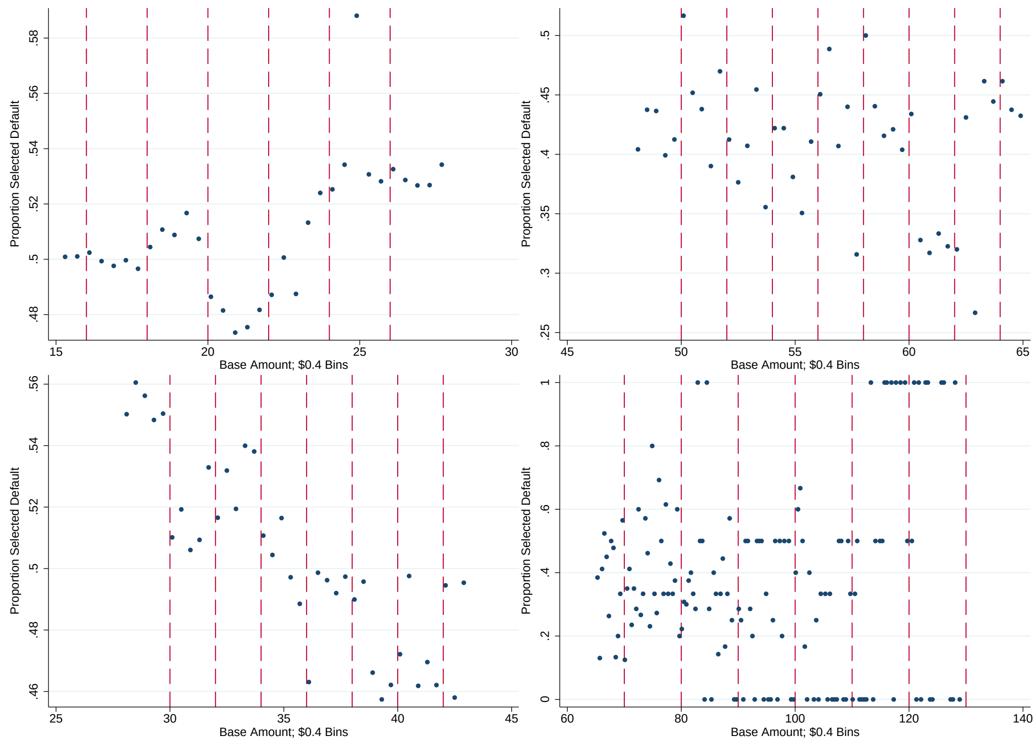
Table B.2: Heterogeneity by Number of Passengers: Regression Discontinuity estimates of Default Effect on Tip Amount

	(1) One b/se	(2) Two b/se	(3) Three b/se	(4) Four b/se
1_[Fare>=15]	0.299***	0.291***	0.267***	0.320***
	(0.006)	(0.012)	(0.024)	(0.048)
Constant	2.401***	2.387***	2.336***	2.386***
	(0.008)	(0.017)	(0.032)	(0.073)
N	3,911,662	851,296	241,251	85,979
r2	0.208	0.225	0.206	0.256

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Robust standard errors clustered at the fare amount relative to the discontinuity ($15). All specifications use 3rd-order polynomials and include fixed effects for driver, pick-up day of the week, pick-up hour, pick-up location borough, and drop-off location borough. The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $5 and less than $25 during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).

Figure B.4: Computational Costs: Proportion of Customers that Use a Default Option by Each $0.40 Fare Bin.



Notes: The sample is limited to rides on Vendor-equipped taxi cabs with fares greater than $15 and lower than $135 and during the first pricing regime (January 1, 2009 - October 31, 2009. 6am - 4pm on Monday - Friday and 6am - 8pm on Saturday and Sunday).